Week 13: All-Pairs Shortest Paths

Algorithms & Data Structures

Thorben Klabunde

th-kl.ch

December 15, 2025

Agenda

1. Quiz
2. Assignment
3. All-Pairs Shortest Paths
4. Floyd-Warshall
5. Johnson
6. Matrix Multiplication
7. Additional Practice
8. Exam Season Primer

Quiz

Q1

Kruskal's algorithm on a connected graph $G = (V, E)$ has a runtime of $O(|E| + |V| \log(|V|))$.

- ○ True
- ◉ False ✓

We need to sort the edges in increasing order at the beginning which takes $O(|E| \log(|E|))$ time, and $|V| \log(|V|)$ does not cover this.

The correct answer is 'False'.

Q2

Suppose we run Kruskal's algorithm on a connected graph $G = (V, E)$. Then for a vertex $u \in V$, consider the quantity N_u defined as the number of times we change $\texttt{repr}[u]$. What is the tightest upper bound for N_u that holds for any connected graph?

a. $N_u \leq O(1)$

b. $N_u \leq O(\log_2(|V|))$ ✓

c. $N_u \leq O(|V|)$

Your answer is correct.

We saw in lecture $N_u \leq \log_2(|V|)$ since we only change it into a representative just as big or bigger, so we will always at least double the size of the representative.

The correct answer is: $N_u \leq O(\log_2(|V|))$

Q3

Asymptotically, Floyd-Warshall's runtime is always at least as fast as Johnson's algorithm's runtime.

- True
- False ✓

No, if say there are only linear (in number of vertices) number of edges, then Johnson's runs faster than Floyd-Warshall.

The correct answer is 'False'.

Q4

After the k-th iteration of the outer loop of the Floyd-Warshall algorithm, $d^k_{u,v}$ contains

a. the length of the shortest $u - v$ path using at most k edges

b. the length of the shortest $u - v$ path that uses only intermediate vertices from $\{1, \ldots, k\}$ ✓

Your answer is correct.

This is how the DP is defined in lecture.

The correct answer is: the length of the shortest $u - v$ path that uses only intermediate vertices from $\{1, \ldots, k\}$

Q5

Suppose you run Floyd-Warshall on a directed, weighted graph $G = (V, E)$ with vertices $V = \{1, 2, 3\}$.

Assume the DP table of d^3 is as follows:

0	-2	-3
2	0	-1
3	1	0

Is it true that any such G will have no negative cycle?

- (●) True ✓
- ( ) False

We saw in lecture that in the final DP table, we have a negative cycle in the graph iff one of the diagonal entries is negative.

The correct answer is 'True'.

Q6

Suppose we have a graph $G = (V, E)$ with weights $c : E \to \mathbb{R}$. Recall in Johnson's algorithm we use Bellman-Ford as a subroutine on a modified version of G to find a function $h : V \to \mathbb{R}$ such that we get a new set of weights $\hat{c} : E \to \mathbb{R}$ defined as $\hat{c}((u, v)) = c((u, v)) + h(u) - h(v)$ for $(u, v) \in E$ which guarantees $\hat{c}((u, v)) \geq 0$.

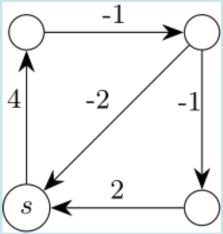
Then we have $h(v) \leq 0$ for all $v \in V$.

- (●) True ✓
- ( ) False

To find h we define a new vertex z and connect it to all vertices with weight zero. Therefore, as $h(v)$ is the distance from z to \(v|), this must always be nonpositive.

The correct answer is 'True'.

Q7



Suppose you run Johnson's algorithm on the above graph and compute the $h : V \to \mathbb{R}$ function as in lecture. What is the value of $h(s)$?

Answer: -3

With the new vertex z, we can see that the path with smallest distance is to go first to the top-left-> top-right -> s, which is distance -3.

The correct answer is: -3

Q8

Let $G = (V, E)$ be a directed graph with vertices $V = \{1, 2, 3, 4\}$.

Let A_G be the *adjacency matrix* of G. Suppose that $(A_G)^3$ (i.e., *the adjacency matrix to the power 3*) is equal to

$$\begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

How many directed triangles (a collection of three edges which form a closed walk) does G have?

Answer: 1

In lecture, we saw that number of triangles is the sum of the diagonal entries divided by 3.

The correct answer is: 1

Q9

Let $G = (V, E)$ be a directed graph, which has self loops on every vertex, and let A_G be its *adjacency matrix*.

Suppose there exists some $i \in \mathbb{N}$ such that A_G^i has all non-zero entries, then all pairs of vertices are reachable from one another.

- (●) True ✓
- ( ) False

A_G^i has a positive entry at index u, v if we can reach u to v with a walk of i steps, so if at some i, all entries are positive, then all vertices are reachable from one another.

The correct answer is 'True'.

Assignment

Remarks on Assignment 12

We still have **quite some ground to cover today** and most exercises relate to the exchange argument and cut property, which we covered in detail in the last two sessions. Therefore, **we will skip a detailed in-class discussion** today.

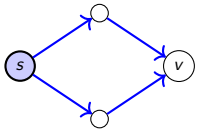
However, the exercises and the introduced notions and terminology (heavy and light edges) are **important**! **Please refer to the Master Solution** and **do not hesitate to reach out if anything should still be unclear!**

All-Pairs Shortest Paths

Recall: Single-Source Shortest Paths (SSSP)

The Problem

Given a source $s \in V$, find the shortest path distances $d(s, v)$ to **all** other vertices $v \in V$.



The Landscape of Algorithms:

Graph Type / Constraint	Algorithm	Runtime
Unweighted	**BFS**	$O(m+n)$ ✓
Weights $c(e) \geq 0$	**Dijkstra**	$O((m+n)\log n)$✓
General Weights	**Bellman-Ford**	$O(nm)$!
DAG (Directed Acyclic)	**Topo Sort + DP**	$O(m+n)$ ✓

Problem Definition: All-Pairs Shortest Paths

Goal

Given directed graph $G = (V, E)$ with edge weights $c : E \to \mathbb{R}$.

Find the shortest path distance $\delta(u, v)$ for **every pair of vertices** $(u, v) \in V \times V$.

Output: An $n \times n$ matrix D where $D_{uv} = \delta(u, v)$.

Edge Cases:

- $\delta(u, v) = \infty$ if v is unreachable from u.
- $\delta(u, v) = -\infty$ if the path touches a negative cycle.

Why do we need this?

Motivation: Pre-computation! Instead of running Dijkstra every time a user asks for a route, we look up the distance in $O(1)$ from our matrix.

Naive Approach: Run SSSP n Times

Idea: Just iterate through every vertex $v \in V$ and run a Single-Source algorithm.

Runtime Analysis for APSP ($n\times$ SSSP):

Weights	**Algorithm**	**Sparse ($m \approx n$)**	**Dense ($m \approx n^2$)**
Unweighted	$n\times$ BFS	$O(n^2)$	$O(n^3)$
Non-negative	$n\times$ Dijkstra	$O(n^2 \log n)$	$O(n^3 \log n)$
General	$n\times$ **Bellman-Ford**	$O(n^2 m) \approx O(n^3)$	$O(n^4)$

The Bottleneck

For general graphs (negative weights allowed), the naive approach is **extremely slow** ($O(n^4)$) on dense graphs!

Question: Can we design a specialized algorithm to beat $O(n^4)$?

Today's Results

Good news: We can beat $n\times$ Bellman-Ford!

New Algorithms

Algorithm	Runtime	Best for
Floyd-Warshall	$O(n^3)$	Dense graphs
Johnson	$O(n(m+n)\log n)$	Sparse graphs

Comparison:

- Floyd-Warshall: Saves factor of n over Bellman-Ford in dense case!
- Johnson: Same as running Dijkstra n times (even with negative weights!)

Counterintuitive: Johnson achieves Dijkstra-like performance with negative weights! How? Because all-pairs is a harder problem, giving us more flexibility...

DP Attempt 1: Subproblem

Obvious subproblems: d_{uv} = shortest path from u to v

Problem: This leads to infinite recursion!

Solution: Add a parameter to make progress

Subproblem Definition (recall Bellman-Ford)

$d_{uv}^{(m)}$ = weight of shortest path from u to v using **at most** m **edges**

- Now we have a natural notion of "smaller": smaller m
- Eventually we solve $d_{uv}^{(n-1)}$ (at most $n-1$ edges)
- If no negative cycles: simple paths have $\leq n-1$ edges, so $d_{uv}^{(n-1)} = \delta(u, v)$

Bonus: Negative cycle detection!
If $d_{vv}^{(n-1)} < 0$ for some v, then negative cycle exists.

DP Method 1: Recurrence

Question: How to compute $d_{uv}^{(m)}$ from smaller subproblems?

Guess: What is the last edge in the shortest path?

Let the last edge be (x, v) for some vertex x

Recurrence

$$d_{uv}^{(m)} = \min_{x \in V} \left\{ d_{ux}^{(m-1)} + c(x, v) \right\}$$

Interpretation: Try all possible last vertices x, take the minimum

Base case:

$$d_{uv}^{(0)} = \begin{cases} 0 & \text{if } u = v \\ \infty & \text{otherwise} \end{cases}$$

Order of computation: Process all $d^{(0)}$, then all $d^{(1)}$, then all $d^{(2)}$, etc.

(Within each m, can compute (u, v) pairs in any order)

DP Method 1: Pseudocode

```
1:  function AllPairsDP1(G = (V, E, c))
2:      // Base case
3:      for u ∈ V do
4:          d_uu ← 0
5:      for u ∈ V, v ∈ V \ {u} do
6:          d_uv ← ∞
7:
8:      // DP iteration
9:      for m = 1 to n − 1 do
10:         for u ∈ V do
11:             for v ∈ V do
12:                 for x ∈ V do
13:                     if d_uv > d_ux + c(x, v) then          ▷ Relaxation step
14:                         d_uv ← d_ux + c(x, v)
15:     return d
```

Runtime: $O(n) \times O(n) \times O(n) \times O(n) = O(n^4)$, not that great!

Floyd-Warshall

Floyd-Warshall: Different Subproblems

Goal: Remove n factor to get $O(n^3)$ by choosing a different subproblem.

Assume vertices are numbered $1, 2, \ldots, n$

New Subproblem (Floyd-Warshall)

d_{uv}^k = weight of shortest path from u to v using only vertices from $\{1, 2, \ldots, k\}$ as **intermediate** vertices

(Note: u and v themselves can be $> k$; only intermediate vertices restricted)

Goal: Compute d_{uv}^n for all pairs (can use all vertices as intermediates)

Progress: Still n^3 subproblems, but...

Floyd-Warshall: The Key Insight

Question: How to compute d^k from d^{k-1}?

Guess: Is vertex k used in the shortest path from u to v?

Two cases:

1. **Vertex k NOT in path:** Then path uses only $\{1, \ldots, k-1\}$
   Cost: d_{uv}^{k-1}
2. **Vertex k IS in path:** Then path is $u \rightsquigarrow k \rightsquigarrow v$
   Cost: $d_{uk}^{k-1} + d_{kv}^{k-1}$

Recurrence:

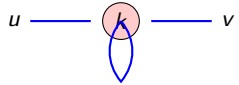
$$d_{uv}^{k} = \min\left\{ d_{uv}^{k-1}, \quad d_{uk}^{k-1} + d_{kv}^{k-1} \right\}$$

Only 2 choices instead of n choices! Constant time per subproblem!

Floyd-Warshall: Why Only 2 Choices?

Before: Guessed which vertex comes last $\rightarrow n$ choices
Now: Guess whether vertex k appears at all $\rightarrow$ 2 choices

What if k appears multiple times?



Creates a cycle through k!

When is this useful?

- If cycle cost > 0: wasteful
- If cycle cost $= 0$: doesn't help
- If cycle cost < 0: infinite loop!

Assumption: No negative cycles
Then using k once is optimal!

Your Turn: Complete Floyd-Warshall

Fill in the blanks to complete the algorithm

```
1:  function FloydWarshall(G = (V, E, c))
2:      // Base case
3:      for u, v ∈ V do
4:          if u = v then
5:              d_uv ← _____
6:          else if (u, v) ∈ E then
7:              d_uv ← _____
8:          else
9:              d_uv ← _____
10:
11:     // DP: gradually allow more intermediate vertices
12:     for k = _____ to _____ do
13:         for u = _____ to _____ do
14:             for v = _____ to _____ do
15:                 if d_uv > _____ then
16:                     d_uv ← _____
17:     return d
```

Floyd-Warshall: The Algorithm

```
1:  function FloydWarshall(G = (V, E, c))
2:      // Base case: k = 0 (no intermediate vertices)
3:      for u, v ∈ V do
4:          if u = v then
5:              d_uv ← 0
6:          else if (u, v) ∈ E then
7:              d_uv ← c(u, v)
8:          else
9:              d_uv ← ∞
10:
11:     // DP: gradually allow more intermediate vertices
12:     for k = 1 to n do
13:         for u = 1 to n do
14:             for v = 1 to n do
15:                 if d_uv > d_uk + d_kv then          ▷ Relaxation
16:                     d_uv ← d_uk + d_kv
17:     return d
```

Runtime: $O(n^3)$

Three nested loops (k, u, v), each runs n times

Work per iteration: $O(1)$

Space:

Naive: $O(n^3)$ (store all d^k matrices)

Optimized: $O(n^2)$ (reuse same matrix, drop superscripts since d^k only depends on d^{k-1})

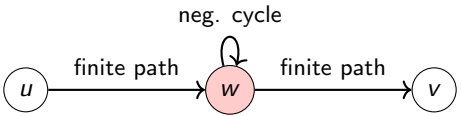
Floyd-Warshall: Handling Negative Cycles

Shortest Path Weight with Negative Cycles

$$\delta(u,v) = \begin{cases} -\infty & \text{if } \exists w : d^n_{uw} < \infty, d^n_{wv} < \infty, \text{ and } d^n_{ww} < 0 \\ d^n_{uv} & \text{otherwise} \end{cases}$$

Interpretation:

- If there exists vertex w on a negative cycle
- AND w is reachable from u (i.e., $d^n_{uw} < \infty$)
- AND v is reachable from w (i.e., $d^n_{wv} < \infty$)
- Then we can loop through the negative cycle infinitely $\rightarrow \delta(u,v) = -\infty$



Johnson

Can We Do Better for Sparse Graphs?

Current best:

- Floyd-Warshall: $O(n^3)$
- For sparse graphs ($m = O(n)$): still $O(n^3)$

Question: Running Dijkstra n times gives $O(n(m+n)\log n)$
For sparse graphs: $O(n^2 \log n)$ — much better than $O(n^3)$!

But... Dijkstra requires non-negative weights.

Crazy idea: What if we could make all edge weights non-negative?
Then we could use Dijkstra even with originally negative weights!

Game-Plan

Johnson's algorithm does exactly this by performing **three steps**:

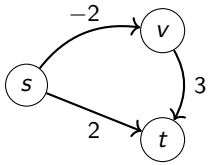
1. **Find a function** $h : V \to \mathbb{R}$ such that $w_h(u, v) = w(u, v) + h(u) - h(v) \geq 0$ for all $u, v \in V$ or determine that a negative-weight cycle exists
2. **Run Dijkstra from every vertex** in the newly weighted graph $G = (V, E, w_h)$ to find $\delta_h(u, v)$
3. **Compute** $\delta(u, v)$ from $\delta_h(u, v)$

We now want to show:

1. *why* this set-up works,
2. *when* we can hope to find such a function h, and
3. *how* we find h

The Naive Attempt at Enforcing Positive Weights Fails

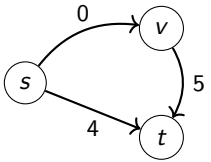
Naive idea: Add constant D to all edges where $D = |\min_{e \in E} c(e)|$



Original costs:

- Via v: $-2 + 3 = 1$ (shortest)
- Direct: 2

Add $D = 2$ to all edges:



New costs:

- Via v: $0 + 5 = 5$
- Direct: 4 (now shortest!)

FAILS! Different paths shifted by different amounts!
2-edge path: $+4$. 1-edge path: $+2$.

Johnson's Insight: Vertex Potentials

Key observation: We need all s-t paths to shift by the **same** amount!

Johnson's reweighting:
Assign each vertex v a "potential" (or "height") $h(v) \in \mathbb{R}$
Define new edge weights: $\hat{c}(u, v) = c(u, v) + h(u) - h(v)$

Claim: For any path $P = (v_0, v_1, \ldots, v_k)$: $\hat{c}(P) = c(P) + h(s) - h(t)$ **Proof:** Let $u, v \in V$ be arbitrary and let P be an arbitrary path from u to v:

$$\begin{aligned}
\hat{c}(P) &= \sum_{i=0}^{k-1} \hat{c}(v_i, v_{i+1}) \overset{\text{Def. of } \hat{c}}{=} \sum_{i=0}^{k-1} \left(c(v_i, v_{i+1}) + h(v_i) - h(v_{i+1})\right) \\
&= \sum_{i=0}^{k-1} c(v_i, v_{i+1}) + \underbrace{\left(h(v_0) - h(v_1)\right) + \left(h(v_1) - h(v_2)\right) + \cdots + \left(h(v_{k-1}) - h(v_k)\right)}_{\text{telescopes to } h(v_0) - h(v_k)} \\
&= c(P) + h(s) - h(t) \qquad \square
\end{aligned}$$

Johnson's Reweighting: The Magic

Consequence of telescoping:

All paths from s to t get shifted by **exactly** $h(s) - h(t)$

$\implies$ Shortest paths are preserved!

Therefore:

- If $\hat{c}(u, v) \geq 0$ for all edges, we can run Dijkstra
- Dijkstra finds shortest paths in $\hat{c}$ graph
- These are also shortest paths in original c graph!
- Just need to convert back: $\delta(u, v) = \hat{\delta}(u, v) - h(u) + h(v)$

Remaining question: How do we find h such that $\hat{c}(u, v) \geq 0$?

Finding h: System of Difference Constraints

Requirement: $\hat{c}(u,v) = c(u,v) + h(u) - h(v) \geq 0$ for all edges (u,v)

Rearrange: $h(v) - h(u) \leq c(u,v)$ for all $(u,v) \in E$

System of Difference Constraints: Find $h : V \to \mathbb{R}$ satisfying:

$$h(v) \leq h(u) + c(u,v) \quad \forall (u,v) \in E$$

We now show that if there are no negative cycles, we can easily find a solution to the system using methods we already know!

Theorem

The system $h(v) - h(u) \leq c(u,v)$ has a solution $\iff$ there exist no negative cycles

Johnson: When does a solution exist?

Theorem: The system $h(v) - h(u) \leq c(u, v)$ has a solution $\iff$ no negative cycles exist

Proof ($\Rightarrow$): By contraposition. Let $C = (v_0, v_1, \ldots, v_k, v_0)$ be a negative cycle and suppose for contradiction that the system was solvable: We get the following system of constraints:

$$\begin{aligned} h(v_1) - h(v_0) &\leq c(v_0, v_1) \\ h(v_2) - h(v_1) &\leq c(v_1, v_2) \\ &\vdots \\ h(v_0) - h(v_k) &\leq c(v_k, v_0) \end{aligned}$$

Adding them up notice that the LHS sums to 0 since for every $v \in C$, $h(v)$ appears as the first (positive) and second (negative) term.

But then: $0 \leq c(C) < 0$, a contradiction, as required. □

Finding h: System of Difference Constraints

($\Rightarrow$) Assume that no negative cycles exist and recall the system we want to solve:

System of Difference Constraints: Find $h : V \to \mathbb{R}$ satisfying:

$$h(v) \leq h(u) + c(u, v) \quad \forall (u, v) \in E$$

Notice that this looks like the **triangle inequality**!

If $h(v) = \delta(s, v)$ for some source s, then triangle inequality guarantees:

$$\delta(s, v) \leq \delta(s, u) + c(u, v)$$

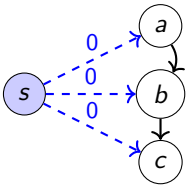
If no negative cycles exist, **shortest path distances satisfy our constraints**!

Johnson's Construction

Problem: Which source s to use? Need to reach all vertices!

Johnson's Trick

1. Add new vertex s to graph
2. Add edges (s, v) with $c(s, v) = 0$ for all $v \in V$
3. Run Bellman-Ford from s
4. Set $h(v) = \delta(s, v)$ for all v



Why it works:

- s can reach all vertices
- No new cycles created
- Triangle inequality holds
- All $h(v) \leq 0$ (paths from s have cost ≤ 0)

Johnson's Algorithm: Putting It Together

1. **Find h:** Add source s, run Bellman-Ford
   - If negative cycle detected: STOP (no solution)
   - Otherwise: $h(v) = \delta(s, v)$ for all v
   - Runtime: $O(nm)$
2. **Reweight edges:** Compute $\hat{c}(u, v) = c(u, v) + h(u) - h(v)$
   - Now all $\hat{c}(u, v) \geq 0$ (triangle inequality!)
   - Runtime: $O(m)$
3. **Run Dijkstra n times:** Once from each vertex
   - Computes $\hat{\delta}(u, v)$ for all pairs
   - Runtime: $n \times O((m + n) \log n) = O(n(m + n) \log n)$
4. **Convert back:** $\delta(u, v) = \hat{\delta}(u, v) - h(u) + h(v)$
   - Runtime: $O(n^2)$

Total runtime: $O(nm + n(m + n) \log n) = O(n(m + n) \log n)$

Matrix Multiplication

Matrix Multiplication Connection

Crazy idea: Shortest paths $\approx$ matrix multiplication.

Standard Matrix Multiplication ($C = A \times B$):

$$C_{ij} = \sum_{k=1}^{n} A_{ik} \times B_{kj}$$

Recall our initial Shortest Path Recurrence: *(attempt 1 before Floyd-Warshall)*

$$d_{ij}^{(m)} = \min_{k=1}^{n} \left\{ d_{ik}^{(m-1)} + w_{kj} \right\}$$

Looking at the indices, these **expressions resemble each other**. But the operations don't match... Or do they?

Choosing the Algebra: Min-Plus

Note: The operations of (naive) matrix multiplication can be performed in **any semiring**.

In particular, we can **define our own semiring algebra** with a **"new" multiplication and addition** operation and apply matrix multiplication.

If the result is meaningful is a different question but we can do so!

	Standard Arithmetic (Field $\mathbb{R}$)	**Min-Plus Semiring** (Tropical Semiring)
Set S	$\mathbb{R}$	$\mathbb{R} \cup \{\infty\}$
Addition ($\oplus$)	$+$	min
Multiplication ($\otimes$)	$\times$	$+$
Add. Identity ($\mathbf{0}$)	0	∞ (since $\min(x, \infty) = x$)
Mult. Identity ($\mathbf{1}$)	1	0 (since $x + 0 = x$)

Matrix Multiplication = APSP

1. The Weight Matrix (W):

W_{kj} = weight of edge $(k \to j)$

2. Distance Matrix ($D^{(m-1)}$):

$D_{ik}^{(m-1)}$ = dist $(i \to k)$ using $\leq m-1$ edges

---

3. The Product Calculation: Let's compute the entry (i, j) of $D^{(m-1)} \otimes W$ (*here* $\otimes$ *as matr. mult*):

$$(D^{(m-1)} \otimes W)_{ij} = \bigoplus_{k=1}^{n} \left(D_{ik}^{(m-1)} \otimes W_{kj} \right)$$

Substitute our Semiring definitions ($\oplus \to \min$, $\otimes \to +$):

$$= \min_{k=1}^{n} \left(\underbrace{D_{ik}^{(m-1)}}_{\text{Path } i \to k} + \underbrace{W_{kj}}_{\text{Edge } k \to j} \right)$$

The Connection

This formula is exactly our DP recurrence: $d_{ij}^{(m)} = \min_k (d_{ik}^{(m-1)} + w_{kj})$

Therefore, one step of Shortest Path is one Matrix Multiplication: $\mathbf{D^{(m)} = D^{(m-1)} \otimes W}$

Concrete Example: Computing One Cell

Let's compute entry $[\mathbf{2}, \mathbf{3}]$ of the product $D \otimes W$.

We need **Row 2** of Left Matrix and **Col 3** of Right Matrix.

$$\begin{pmatrix} \cdot & \cdot & \cdot \\ \infty & \mathbf{0} & \mathbf{2} \\ \cdot & \cdot & \cdot \end{pmatrix} \otimes \begin{pmatrix} \cdot & \cdot & \mathbf{8} \\ \cdot & \cdot & \mathbf{1} \\ \cdot & \cdot & \mathbf{0} \end{pmatrix}$$

Abstract (Semiring)

Formula: $\bigoplus(Row_k \otimes Col_k)$

$k = 1$: $\infty \otimes 8$

$k = 2$: $\oplus \quad (0 \otimes 1)$

$k = 3$: $\oplus \quad (2 \otimes 0)$

Result: $(\infty \otimes 8) \oplus (0 \otimes 1) \oplus (2 \otimes 0)$

Concrete (Intuition)

Meaning: $\min(Row_k + Col_k)$

$k = 1$: $\infty + 8 = \infty$

$k = 2$: $\min(0 + 1) = 1$

$k = 3$: $\min(2 + 0) = 2$

Result: $\min(\infty, 1, 2) = \mathbf{1}$

The shortest path from node 2 to 3 going through an intermediate node k is length 1.

APSP via Iterative Squaring

Notice: To compute APSP, we simply need to compute the n-th power! We can do so efficiently using iterative squaring:

$$
\begin{aligned}
W^{\otimes 1} &= W \\
W^{\otimes 2} &= W^{\otimes 1} \otimes W^{\otimes 1} \\
W^{\otimes 4} &= W^{\otimes 2} \otimes W^{\otimes 2} \\
W^{\otimes 8} &= W^{\otimes 4} \otimes W^{\otimes 4} \\
&\vdots
\end{aligned}
$$

Only $\log n$ **multiplications** needed to reach $W^{\otimes n}$!

Runtime: $O(\log n) \times O(n^3) = O(n^3 \log n)$. Not yet better than Floyd-Warshall...

Can we use one of the more efficient matrix multiplication algorithms to beat Floyd-Warshall? For instance, **Strassen's algorithm**?

Why Not Faster Matrix Multiplication?

Bad news: Strassen ($O(n^{2.807})$) requires subtraction but min has no inverse:

- Given $\min(a, b) = c$, cannot recover a and b
- The min-plus structure is a **semiring**, not a ring
- No "minus" operation exists

Conclusion: For all-pair shortest paths, we do not know a better way to perform matrix multiplication than in $O(n^3)$ yielding $O(n^3 \log n)$ through iterative squaring.

However, there IS a related problem where fast matrix multiplication helps!

All-Pair Reachability (Transitive Closure)

Transitive Closure Problem

Input: Directed graph $G = (V, E)$ with adjacency matrix A. Output: For all pairs (i, j), does there exist **any** path from i to j?

$$T[i, j] = 1 \iff \exists \text{ path } i \rightsquigarrow j$$

Algebra: Boolean Semiring

- **Set:** $\{0, 1\}$
- **Addition ($\oplus$):** $\vee$ (OR) $\rightarrow$ "Is there a path via neighbor 1 OR 2?"
- **Multiplication ($\otimes$):** $\wedge$ (AND) $\rightarrow$ "Step to k AND then $k \rightarrow j$?"

Note: Self-Loop Trick

Add Self-Loops to Allow Waiting

Idea: Add a self-loop to every node ($A' = A + I$).

- Meaning: You can "wait" at a node for a step.
- A path of length 1 ($u \to v$) becomes a path of length 2 ($u \to v \to v$).

Result: $(A + I)^{n-1}$ captures all paths of length $\leq n - 1$.

Example: Graph $1 \to 2 \to 3$. Can 1 reach 2 in ≤ 2 steps?

Without Self-Loops (A^2)

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \to A^2 = \begin{pmatrix} 0 & \mathbf{0} & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Missed $1 \to 2$! It found only length exactly 2 ($1 \to 3$).

With Loops ($(A + I)^2$)

$$A' = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \to (A')^2 = \begin{pmatrix} 1 & \mathbf{1} & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Found $1 \to 2$! Logic: $1 \to 2 \to \mathbf{2}$ (Wait).

Can we use Fast Matrix Multiplication?

Standard algorithms like Strassen ($O(n^{2.81})$) require a **Ring** (subtraction).

The Workaround:

1. Treat Boolean matrices as **Integer** matrices ($0, 1 \in \mathbb{Z}$).
2. Compute product using Fast MatMul over Integers.
3. Map any result > 0 to 1 (Boolean TRUE).

Result: Transitive Closure is solvable in roughly $O(n^{2.37})$ using the fastest known matrix multiplication algorithm. This is the fastest algorithm we know to solve the transitive closure!

Path Counting (Standard Arithmetic)

The Question

How many distinct walks of **exactly length** k exist from node i to node j?

The Algebra: Use standard arithmetic!

- $\oplus \to +$ (Addition sums up the options)
- $\otimes \to \times$ (Multiplication combines steps in a sequence)

Theorem

If A is the adjacency matrix (0 or 1), then:

$$(A^k)_{ij} = \text{Number of walks from } i \text{ to } j \text{ with length } k$$

Application (Triangle Counting): The number of triangles in a graph is $\text{Trace}(A^3)/3$.

- A^3_{ii} counts paths $i \to \cdots \to \cdots \to i$ (cycles of length 3).
- Divide by 3 because each triangle is counted once for each vertex.

Questions?

Additional Practice

Theory Task T4. / 12 P

Assume that there are n towns $T_1, \ldots, T_n$ in the country Examistan. For each pair of distinct towns T_i and T_j, there is exactly one road from T_i to T_j. All of the roads in Examistan are one-way. This implies that there is always a road from T_i to T_j and another road from T_j to T_i. Each road has a nonnegative integer cost that you need to pay to use this road.

For simplicity you can assume that each town T_i is represented by its index i.

In the following subtasks b) and c), you can assume that the directed graph in a) is represented by a data structure that allows you to traverse the direct successors and direct predecessors of a vertex u in time $\mathcal{O}(\deg_+(u))$ and $\mathcal{O}(\deg_-(u))$ respectively, where $\deg_-(u)$ is the in-degree of vertex u and $\deg_+(u)$ is the out-degree of vertex u.

/ 6 P

b) Due to the epidemiological situation in Examistan, the authorities decided to reduce the number of trips between different towns. Now the only way to get from one town to another is to use the roads. Moreover, if you want to travel from town T_i to the other town T_j, you must visit a test center during your trip (in T_i or T_j or elsewhere with a detour). Since test centers are expensive, there are only $k < n$ of them, and they are located only in the first k towns $T_1, \ldots, T_k$ (i.e., one test center in each of these towns).

Assume that you need to fill the table of minimal costs required to travel between all pairs of towns, which takes into account the new rules of travelling. Provide an as efficient as possible algorithm that takes as input a graph G from task a) and a number k, and outputs a table C such that $C[i][j]$ is the minimal total cost of roads that one can use to get from T_i to T_j while also visiting a test center. You can assume that for all $1 \leq i \leq n$, $C[i][i] = 0$.

What is the running time of your algorithm in concise Θ-notation in terms of n and k? Justify your answer.

FS20 Solution

Solution: The towns are modeled as the vertices $V = \{1, \ldots, n\}$ of the graph G. The roads are modeled as directed edges $E = \{(i, j) \mid i \neq j, i, j = 1, \ldots, n\}$. The costs that you need to pay to use the roads are modeled as the weights w of the respective edges.

The number of vertices is thus $|V| = n$ and the number of edges $|E| = n^2 - n$, since n^2 is the number of possible ordered pairs (i, j) and we have to subtract the n self-edges represented by (i, i) as they are not part of our graph.

Alternative way to get the number of edges: you choose 2 out of n to get the number of unordered sets $\{i, j\}$ with $i \neq j$, resulting in $\binom{n}{2} = \frac{1}{2}(n-1)n$. But we care for the different directions so we have to multiply this number by 2 (for (i, j) and (j, i)) resulting in $|E| = n^2 - n$.

FS20 Solution

function ModifiedFloydWarshall(G)
  $C_{\text{from}}, C_{\text{to}}$ ▷ Cheapest paths from/to $T_1, \ldots, T_k$, initial infinity
  for ℓ from 1 to n **do**
    for i from 1 to n **do**
      for j from 1 to k **do**
        if $C_{\text{from}}[i][j] > w_{i,\ell} + w_{\ell,j}$ **then**
          $C_{\text{from}}[i][j] \leftarrow w_{i,\ell} + w_{\ell,j}$
        if $C_{\text{to}}[j][i] > w_{\ell,i} + w_{j,\ell}$ **then**
          $C_{\text{to}}[j][i] \leftarrow w_{\ell,i} + w_{j,\ell}$
  $C[i][i] \leftarrow 0$ ▷ Path to itself is zero for each node
  for ℓ from 1 to k **do**
    for i from 1 to n **do**
      for j from 1 to n **do**
        if $C[i][j] > C_{\text{from}}[i][\ell] + C_{\text{to}}[\ell][j]$ **then**
          $C[i][j] \leftarrow C_{\text{from}}[i][\ell] + C_{\text{to}}[\ell][j]$

The algorithm is basically two times Floyd-Warshall, hence we obtain a running time of $\Theta(n^2k)$, as in our modified version one loop only goes until k and not n.

Exam Season Primer

You Made It!

Congratulations on reaching the end of the term!

The toughest leg is behind you.

Now you have:

- No new material
- One month of quiet time
- Time to consolidate everything you've learned

Feeling overwhelmed? Need to catch up? **That's completely normal!**

Don't *underestimate* your time
You have a whole month just for yourself!

Don't *overestimate* it either
Stay structured and disciplined!

You can do this!

My Study Strategy: Overview

Key principles:

- Work in **exam order** (first exam first)
- **Frontload** old exams (your best prep!)
- Keep a **rotation** across subjects
- Track your **mistakes** systematically

Three-phase approach for each subject:

1. **Broad overview** (3-4 days)
2. **Old exam kick-off** (one old exam per day for 3-4 days)
3. **Rotation mode** (one old exam every 2 days + review theory review sessions ongoing)
4. **Sprint** (week before exam)

Important: This is *my* system. Adapt it to what works for you!

Phase 1: Broad Overview (3-4 days)

Goal: Refresh memory, identify gaps, feel comfortable again

What to do:

- **Go through** all **your notes** on theory
- **Patch up** anything you don't remember well
- **Solve all old assignments** again
- Start writing your **cheatsheet** (*if applicable*)

Expected outcome: is NOT that you know everything perfectly but instead:

- Feel comfortable, have an overview
- See that everything is manageable

Phase 2: Old Exam Kick-Off (3-4 days)

Old exams are by far the best prep!

My approach:

- Solve one exam per day (3-4 days), while phasing in next subject
- **Always under timed conditions!**
- Properly correct and understand mistakes

Example schedule:

- 8:00-11:00 — Write exam (timed!)
- 11:00-12:00 — Corrections
- 12:00-12:45 — Lunch break
- 13:00-14:30 — Post-exam deep dive (if needed)
- **Rest of the day - Next subject**

Mistakes show you exactly where to focus

My mistake-tracking system:

1. For every subject, keep a list of all mistakes in quiz format
   - Question you got wrong
   - Correct answer
   - Why you made the mistake
2. Review this list periodically (I try to do it every night before bed)
3. Use it to identify areas needing more practice

Why this works:

- Shows specific gaps in understanding
- You still have time to address them
- Periodic revision = long-term retention
- Focuses your efforts where needed most

Phase 3: Rotation Mode

After initial review and first set of old exams, switch to half-day every two days (old exams and some theory review sessions.

New **slots open up** and **rotate in next subject** following the same principle.

You end up in a rotation of half-days of every subject and should eventually mainly focus on solving old exams.

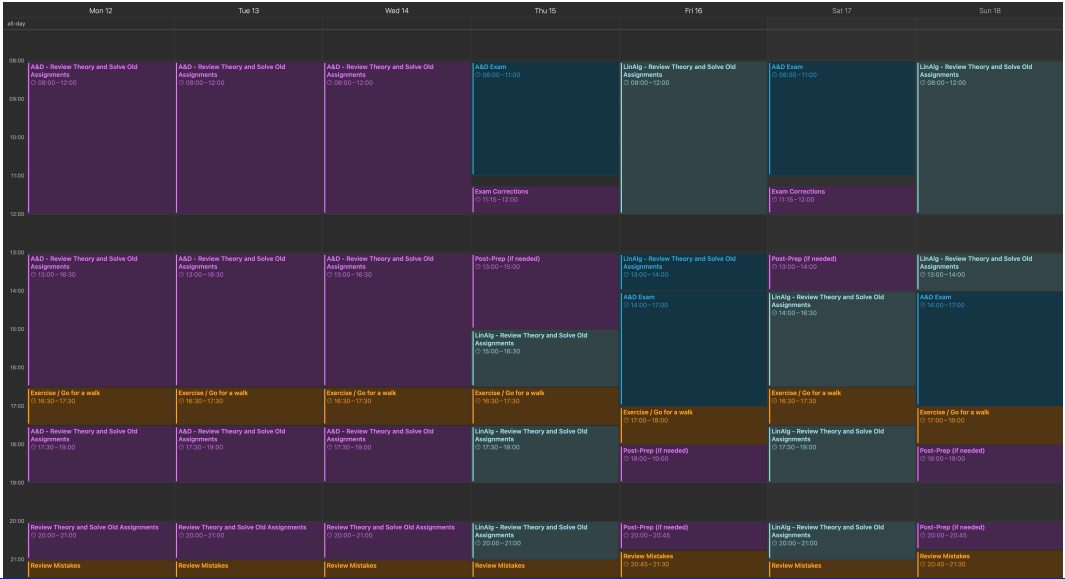
Adjust based on your needs:

- Struggling with one subject? Give it two half-days
- Confident with another? One session every 3 days

Why I like this system:

- Clear structure (no brain power wasted on "what to study?")
- Prioritizes by exam schedule
- Ensures you keep track of what you don't know
- Regular revision of all subjects

Example


Mon 12
Tue 13
Wed 14
Thu 15
Fri 16
Sat 17
Sun 18
all-day
08:00
09:00
10:00
11:00
12:00
13:00
14:00
15:00
16:00
17:00
18:00
19:00
20:00
21:00
A&D - Review Theory and Solve Old Assignments
08:00–12:00
A&D Exam
08:00–11:00
Exam Corrections
11:15–12:00
LinAlg - Review Theory and Solve Old Assignments
08:00–12:00
A&D - Review Theory and Solve Old Assignments
13:00–16:30
Post-Prep (if needed)
12:00–15:30
LinAlg - Review Theory and Solve Old Assignments
13:00–14:00
A&D Exam
14:00–17:00
LinAlg - Review Theory and Solve Old Assignments
14:00–16:30
Post-Prep (if needed)
13:00–16:00
LinAlg - Review Theory and Solve Old Assignments
13:00–14:00
A&D Exam
14:00–17:00
LinAlg - Review Theory and Solve Old Assignments
16:30–17:30
Exercise / Go for a walk
16:30–17:30
A&D - Review Theory and Solve Old Assignments
17:30–19:00
Exercise / Go for a walk
17:00–18:00
Post-Prep (if needed)
18:00–19:00
Review Theory and Solve Old Assignments
20:00–21:00
Review Mistakes
20:45–21:30
Post-Prep (if needed)
20:00–20:45

Study Effectively, Not Just Long

Ineffective: Skim-reading summaries repeatedly

Effective: Active engagement

- Explain concepts out loud
- Compute, derive, prove
- Solve exam questions
- Test yourself without looking at notes

Periodic theory revision:

- Don't *only* do old exams
- Schedule theory review sessions
- Mark areas you don't remember well
- Focus on those in next session

Take Care of Yourself!

Don't Fall Into These Traps

- Cutting back on sleep
- Skipping exercise
- Isolating yourself completely

Sleep:

- Aim for 8 hours every night
- You **don't** need to **sacrifice sleep**
- It will come back to haunt you

Social:

- **Keep in touch** with friends
- **Support** each other, there's **no competition**!

Exercise:

- **Quick sessions do wonders!**
- 20min run or workout
- 10min walks after meals
- Don't go all out, just get heart rate up

You're Not Alone — Resources Available

This is your first exam season, don't be too hard on yourself! Finding a structure that works for you also takes some time, that's normal.

Adapt to what works for you!

If you ever feel too stressed out or notice that you're burning yourself out, act early and think about how to make it more sustainable for you.

Also, **there's no shame in asking for help!**

Available resources:

- **Nightline** — Peer listening service
- **Friends & Study Groups** — You're in this together!
- **Me!** — Feel free to reach out with questions or concerns

Keep Perspective

It's never the end of the world if something doesn't go to plan

Important reminders:

- A bad exam $\neq$ failure
- Grading is **curved** (different from high school!)
- Scale determined *after* exam based on everyone's performance
- Once an exam is over, forget it and move on
- Focus on the next one

You Got This!

Good luck, you're gonna do great!